DOCUMENT RESUME

ED 209 631 CS 006 319

AUTHOR Vande Kopple, William J.
TITLE Experiments on the Readability of Natural Expository Paragraphs with Identical or Related Sentence Topics.
PUB DATE [81]
NOTE 15p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Coherence; Cohesion (Written Composition); College Students; *Connected Discourse; *Discourse Analysis; Expository Writing; Higher Education; High Schools; High School Students; *Paragraph Composition; *Readability; *Reading Research; Writing Research

ABSTRACT

To test the hypothesis that paragraphs composed of sentences with identical or closely related topics (the grammatical subject and its adjuncts) would be easier to read than a paragraph whose sentence topics were only remotely related, two experiments on the readability of paragraphs were conducted. The first experiment involved subjective judgments by 40 high school and 20 college students on the readability of two pairs of paragraphs. The second experiment involved five-minute timed typing tests of pairs of paragraphs. Subjects participating in this second experiment were 14 beginning typing students, 18 students with 36 weeks of typing instruction, and 12 secretarial students with 36 weeks of typing instruction. The data from these experiments strongly supported the hypothesis. The students in the first experiment tended to choose paragraphs with identical or related topics, and the typists tended to make fewer errors while typing at a faster rate using the paragraph with identical or related topics. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED209631

Experiments on the Readability of Natural Expository Paragraphs with Identical or Related Sentence Topics

William J. Vande Kopple

Calvin College

I designed the two experiments described here to test whether a natural expository paragraph composed of sentences with topics that are identical or closely related to each other is easier to read than a paragraph similar or identical to the first in truth value and all other important respects except that the topics of its sentences are related to each other only remotely. In most English sentences, the topic usually includes the grammatical subject and its adjuncts and expresses information that is given in or derivable from the preceding sentences.

My first experiment involved subjective judgments on the readability of pairs of paragraphs. I wrote two pairs of paragraphs. (The four paragraphs are reproduced in the appendix.) Paragraphs 1a and 1b both dealt with some effects of the drug sotolol on hypertensive patients. Paragraphs 2a and 2b both dealt with some effects of the drug pindolol on hypertensives. The sentence topics in 1a and 2a were identical or very similar within the appropriate paragraph. Those in 1b and 2b were only remotely related within the appropriate paragraph. In 1a there were nine sentence topics identical or closely related to each other, one topic related to the information in the comment (The comment of most English sentences includes the verb and objects.) of the sentence immediately before it, and just one topic only remotely related to previous topics or comments. Similarly, in 2a there were eight sentence topics identical or closely related to each other, two topics related to

CS006319

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
William J. Vande Kopple

the information in the sentence comment immediately before them, and no topics only remotely related to previous topics or comments. On the other hand, 1b had only two topics identical or closely related to each other, no topics related to the information in the immediately preceding sentence comment, and seven topics remotely related to previous topics or comments. Likewise, 2b had only one topic closely related to a previous topic, only one related to the information in the sentence comment just before it, and eight remotely related to previous topics or comments.

In many other important respects 1a and 1b were identical. Both had 111 total words and seven orthographic sentences; thus both averaged 15.86 words per sentence. And both had seven main clauses, one adjective clause, and the same orienters (Orienters are introductory words or phrases that orient readers to the content of the sentences they introduce; for example, "in many ways," and "to a certain extent" are orienters.). In several other important ways, 1a and 1b were very similar. 1a contained four adverb clauses while 1b had three. 1a had four passive verbs and four nominalizations to three passives and seven nominalizations in 1b. Also, four grammatical subjects in 1a and five grammatical subjects in 1b did not coincide with the agents of the actions underlying their clauses. Finally, 1a contained ten medical terms while 1b had nine.

In many of these ways, 2a and 2b were also identical. Both had 110 total words and seven orthographic sentences; thus their sentences averaged 15.71 words. Both had seven main clauses, all the same orienters, three passive verbs, four nominalizations, and twelve medical terms. In several other ways, 2a and 2b were very

similar. 2a contained two adjective and two adverb clauses, while 2b contained three of each type of clause. And three grammatical subjects in 2a and four in 2b did not coincide with the agents of the actions underlying their clauses.

Once I had written the four paragraphs, I typed each one with the same IBM typewriter on a sheet of standard typing paper in lines about 70 spaces long. Then I made as many copies of these on the same Xerox copier as I needed. I wrote two different sets of instructions. From one, certain subjects would learn that they would be reading two paragraphs of expository prose dealing with subject matters that were nearly identical, that they should read each paragraph carefully but just once, and that they should indicate whether the paragraph they read first was easier to read and follow than the one they read second, whether the one they read second was easier than the one they read first, or whether they really could not detect any significant difference in ease of reading between the two paragraphs. The other set of directions was the same except that it indicated that those who saw it would be reading two paragraphs on somewhat different subject matters.

I used one of these sets to direct subjects in one treatment and the other to direct subjects in another treatment in this experiment. In the first treatment, subjects received the first set of directions followed by two sheets on which appeared either 1a on the first and 1b on the second, 1b on the first and 1a on the second, 2a on the first and 2b on the second, or 2b on the first and 2a on the second. In the second treatment, subjects received the second set of directions followed by two sheets on which appeared either 1a on the first and 2b on the second, 2b on the

first and 1a on the second, 2a on the first and 1b on the second, or 1b on the first and 2a on the second.

Two different populations participated in each treatment. Forty high-school juniors and seniors and twenty college sophomores had to judge either 1a and 1b or 2a and 2b. Thirty-one high-school juniors and seniors and twenty-four college sophomores had to judge either 1a and 2b or 2a and 1b.

Although I had some misgivings about the second treatment because the versions of paragraph 1 were not as similar to those of paragraph 2 as I would have liked (they should have been nearly identical in syntax, word-forms, and word recognizability), my hypothesis was that paragraphs 1a and 2a, those with identical or closely related sentence topics, would be chosen as the easier of whatever pair they appeared in.

The results in the first treatment confirmed my hypothesis quite strongly. Of the forty high-school juniors and seniors who had to decide between either 1a and 1b or 2a and 2b, ten chose 1a as easier than 1b, another ten chose 1b over 1a, fifteen chose 2a over 2b, and five chose 2b over 2a. Because of the large number who favored 2a over 2b, twenty-five high-school juniors and seniors preferred an "a" paragraph, and only fifteen preferred a "b." Of the twenty college sophomores, five chose 1a as easier than 1b, three chose 1b over 1a, ten chose 2a over 2b, and two chose 2b over 2a. Thus fifteen preferred an "a" paragraph, and five preferred a "b." Summing these results, we find that fifteen students favored 1a over 1b, thirteen favored 1b over 1a, twenty-five favored 2a over 2b, and seven favored 2b over 2a. In all, forty students preferred an "a" paragraph, one with identical

or related sentence topics, and exactly half that many preferred a "b" paragraph, one with unrelated sentence topics.

The evidence on paragraphs 2a and 2b is quite strong. And although the evidence on 1a and 1b is weaker, it does support my hypothesis. One problem that became evident, however, was that thirty-two of the forty high-school students and sixteen of the twenty college students said that the second of the two paragraphs they had read was the easier. Evidently, encountering the same or similar material in a second paragraph probably led some students to believe that that paragraph was easier even though it might have been stylistically more difficult. One way to compensate for this in future experiments would be to have all subjects read the two paragraphs as often as they wished before deciding which of the two seemed easier to read.

Another way would be to proceed as I did in the second treatment of this experiment, that is, to pair a paragraph on one subject matter and with sentences having identical or related topics with a paragraph on another subject matter and with sentences having unrelated topics. As I have noted, however, 1a and 2b and 2a and 1b were not as similar as they should have been; they did not differ just in subject matter and sentence topicality. Thus we must view the results of the treatment that involved them with some caution. I include them only to report results in what I view as a preliminary step toward a more controlled experiment.

In this treatment, the students' responses favored the paragraphs with identical or related sentence topics slightly. Of the thirty-one high-school juniors and seniors who had to judge either 1a and 2b or 1b and 2a, seven chose 1a over 2b, ten chose 2b over

1a, eleven chose 2a over 1b, and three chose 1b over 2a. True, those who chose 2b over 1a outnumbered those who chose 1a over 2b. It is possible that 1a was harder than 2b in ways that are unrelated to sentence topics. But because of the figures for the other pair, a total of eighteen juniors and seniors preferred an "a" paragraph, one with identical or related sentence topics, as contrasted to thirteen who preferred a "b" paragraph, one with unrelated sentence topics. Of the twenty-four college sophomores who received this treatment, four chose 1a as easier than 2b, eight chose 2b over 1a, nine chose 2a over 1b, and three chose 1b over 2a. Again, 2b emerged as easier than 1a, but because of the numbers for the other pair, thirteen college sophomores preferred an "a" paragraph while eleven preferred a "b."

It is debatable how valid an indicator of readability such a subjective judgment really is. Yet probably no one would deny that it has some validity nor that experiments employing it, along with those employing subjective judgments following numerous readings of two paragraphs, can serve as excellent preliminaries to more valid and controlled readability research.

The second experiment I performed involved what I view as a more reliable indicator of readability. It also provided much more persuasive evidence in support of my hypothesis.

For this experiment, one involving five-minute timed typing tests, I wrote one pair of paragraphs, 3a and 3b. (3a and 3b are reproduced in the appendix.) Each paragraph dealt with certain aspects of the construction and performance of a brand of snow skis called the Hart Queen. The major difference between the two paragraphs was in the nature of their sentence topics. Each paragraph

had eighteen sentence topics, but 3a contained only one topic that was only remotely related to previous topics or comments. All the others were identical or closely related to each other or to the information in the sentence comment just before them. On the other hand, 3b contained fourteen topics that were not identical or closely related to previous topics or comments.

In many other ways, 3a and 3b were identical. Both had seventeen sentences, eighteen main clauses, one noun clause, and three adverb clauses. Looking yet more closely at their sentences, I found five pairs of sentences corresponding in position and subject matter that were composed of exactly the same words. Of course, the ordering of the words in the sentences in 3a was different from that in 3b. Moreover, I found six pairs of corresponding sentences that were equally long. And nine of the full verbs in 3a were exactly the same as the corresponding full verbs in 3b. Finally, both paragraphs had the same orienters, two passives, three nominalizations, and two subjects that did not coincide with the agents of the actions underlying their clauses.

In several other ways, 3a and 3b were very similar. 3a had 252 words; thus its orthographic sentences averaged 14.823 words. 3b had 249 words; its orthographic sentences averaged 14.647 words. There was no sentence in 3a that differed from its correspondent in 3b by more than four words. And five of the full verb forms in 3a differed from their corresponding verb forms in 3b only because of changes in helping verbs or inflections necessitated by different numbers or tenses.

Once I had written these two paragraphs, I counted how many

individual letters each contained just as a typing instructor would do. 3a had 1492, while 3b had 1474. I typed each with the same IBM typewriter on a sheet of standard typing paper in lines about 70 spaces long. I made as many copies of these on the same Xerox copier as I needed. Then I prepared a sheet of instructions for my forty-four subjects, students in three typing courses in a local high school. On this sheet, I asked each student to write his or her name and the name of the appropriate course and to prepare for a five-minute typing test. At the start of the five minutes in the actual experiment, half of the subjects flipped the instruction sheet over, saw paragraph 3a, and typed, with as much accuracy as possible, as much of it as they could. The other half typed 3b. Exactly one week after each subject had taken this five-minute test, the half that had typed 3a typed 3b, and the half that had typed 3b typed 3a.

When I received the completed sets of tests, I compared the number of letters of each paragraph that every subject typed. I assumed that the paragraph a subject typed faster was the one he found easier to process. I also compared the number of errors each made while typing 3a and 3b. These comparisons showed that 3a must have been easier to process and to type.

Fourteen of my subjects were in a Typing I class, a class that had had about eighteen weeks of typing experience. Ten of these subjects typed 3a faster than 3b. Of these ten, six typed 3a before 3b. They averaged 41.167 more letters on 3a and made an average of 7.67 errors on 3a as compared to an average of 6.5 errors on 3b. Four other students typed 3a faster than 3b but typed it after 3b. They averaged 22.25 more letters on 3a, while

making an average of 7.25 mistakes. On 3b they averaged 12 errors.

Four of the fourteen students in this class typed 3b faster than 3a. Three of them typed 3b before 3a and averaged forty-four more letters on 3b while making an average of thirteen mistakes on 3b and an average of 8.334 mistakes on 3a. The single subject who typed 3b faster than and after 3a typed two more letters on 3b and made sixteen mistakes on it as compared with fifteen mistakes on 3a.

Eighteen other subjects were in a Typing II class, a class that had had about thirty-six weeks of typing experience. Thirteen of them typed 3a faster than 3b. Of these thirteen, three typed 3a before 3b. They averaged 130.334 more letters on 3a and made an average of twelve errors on it as compared with an average of 13.334 errors on 3b. Ten other students typed 3a faster than and after 3b. These ten averaged 101.9 more letters on it. Their average numbers of errors for 3a and 3b respectively were 9.5 and 12.6.

The other five students in this class typed 3b faster than 3a. All of them typed 3b after 3a. They averaged 43.8 more letters on 3b and made an average of 10.6 mistakes on it. On 3a they averaged 13.8 mistakes.

The final twelve of my subjects were students in a Secretarial Practice course. Before they began this course, they had received 36 weeks of typing instruction. Once in the class, they had typed often but not exclusively for another 36 weeks. Eleven of these students typed 3a faster than 3b. Five of these eleven typed 3a before 3b. They averaged 163.8 more letters on 3a and made an average of eleven mistakes on it as compared to an average of 11.2

on 3b. Six others typed 3a faster than and after 3b. They averaged eighty-five more letters on 3a, twelve errors on 3a, and 15.17 errors on 3b.

The single student from this class who typed 3b faster than 3a did it after she had typed 3a. She typed 197 more letters on 3b and made five mistakes while doing so. On 3a she made nine mistakes.

Since the more highly trained typists who excelled on 3a had greater differences between their scores for 3a and 3b than did the less highly trained typists who excelled on 3a, it would be interesting and valuable to have typists better than the best ones in this experiment type the two paragraphs. To a point, the differences between their scores would probably increase. But it is possible that these differences might decrease as the typists approached professional excellence. It is probable that professional typists linguistically process what they type very little.

Summing the figures for all subjects, we find that thirty-four of forty-four typed 3a faster than 3b. Fourteen of these typed 3a before 3b, they averaged 110.857 more letters on 3a, and they made an average of 9.79 mistakes on 3a as compared to an average of 9.64 mistakes on 3b. The difference between the two mistake averages is so slight that it is really not justified to claim that these fourteen students excelled on 3a at the expense of more errors. The remaining twenty of the thirty-four typed 3a after 3b. They averaged eighty-one more letters on 3a and made an average of 9.8 mistakes while typing it. On 3b they averaged 13.25 errors. Obviously, no one could claim that these twenty excelled on 3a at the expense of more errors.

However, some researchers could maintain that, since twenty of the thirty-four who excelled on 3a typed it after 3b, they did so because they were familiar with its content. Thus they could argue that the total of thirty-four is not quite as significant as it might at first appear. This criticism is probably valid, and the best way to avoid it in future experiments would be to have all subjects type 3a before 3b. If they still were to average more letters on 3a, it would very likely not be because they were familiar with its specific content.

Ten of my subjects typed 3b faster than 3a. Three of these typed 3b before 3a, averaging forty-four more letters and thirteen errors on 3b. On 3a they averaged 8.334 errors. These three might indeed have excelled on 3b at the expense of more errors. The other seven typed 3b after 3a, averaging 59.714 more letters and 10.57 errors. On 3a they averaged 13.29 errors. Obviously, these seven did not excel on 3b at the expense of more errors. Yet we can see that most of those who did better on 3b, in this case more than two-thirds, might have done so because of familiarity with its content.

These data strongly support my hypothesis that a paragraph composed of sentences with topics that are identical or closely related to previous topics or to the information in the sentence comment just before them is easier to read than one identical or similar to it in truth value and in all other important ways except that its sentence topics are only remotely related to each other or to the information in the comment just before them. Additional experimentation with greater numbers of subjects in tests with additional controls should reinforce these findings.

APPENDIX

Paragraph 1a (with its main topics underlined):

Sotolol must be examined further before it can be considered a reliable anti-hypertensive drug. In one test, it caused systolic pressures in resting white mice to fluctuate after they received ten-milligram doses at five-minute intervals. In another, it decreased systolic pressures in standing males. However, in other studies it has led to ambiguous results. For example, when sotolol was used in 50-milligram doses on standing females, it did not decrease their systolic pressures. Yet when it was combined with alprenol it did decrease their pressures for four hours. Finally and most importantly, sotolol might cause fibrosing mediastinitis in people who have serum cholesterol levels 50 milligrams above the national level.

Paragraph 1b (with its main topics underlined):

Further studies must be made on sotolol before its reliability as an anti-hypertensive drug is established. In one test, fluctuations in systolic pressures in resting white mice resulted after they received ten-milligram doses at five-minute intervals. In another, standing males experienced a decrease in systolic pressures because of it. However, ambiguous results have been evident in other studies. For example, systolic pressures in standing females did not decrease when 50-milligram doses of sotolol were used. Yet for four hours a combination of it and alprenol decreased their pressures. Finally and most importantly, fibrosing mediastinitis might result from it in people who have serum cholesterol levels 50 milligrams above the national level.

Paragraph 2a (with its main topics underlined):

Pindolol is useful in treating patients only in special cases. If the doses of pindolol are below ten milligrams, it can aid the blood circulation in almost all chronic heart patients. Also, when combined with a high-protein diet, it often reduces standing diastolic pressures in almost all hypertensives. But pindolol should not be used with patients who suffer from some form of renal failure. When it is given to them, it raises their plasma renin levels dangerously. Nor should pindolol be used on people who have had rheumatic fever at some time in their lives. In them it leads both to extreme dizziness and to irregularities in the nervous system.

Paragraph 2b (with its main topics underlined):

Only in special cases do patients benefit from pindolol. Blood circulation in almost all chronic heart patients is increased if the doses of pindolol are below ten milligrams. In almost all hypertensives, standing diastolic pressures decrease when the treatment includes pindolol combined with a high-protein diet. But patients who suffer from some form of renal failure should not be treated with pindolol. Their plasma renin levels rise dangerously

when it is given to them. People who have had rheumatic fever at some time in their lives are others on whom pindolol should not be used. Extreme dizziness and irregularities in the nervous system are the effects of it in them.

Paragraph 3a (with its main topics underlined):

The Hart Queen is an excellent ski for beginning and intermediate skiers. Its core is a very thin layer of tempered ash, ash direct from the hardwood forests of Kentucky. And its outer construction involves two major innovations to provide for added strength and flexibility. For increased strength, it has two sheets of ten-gauge steel molded securely to its layer of ash. For increased flexibility, it has a wrapping of highly active fiberglass around its two steel sheets. As a result, the Queen's flexibility ratio is an impressive three to one. It develops an even better ratio on difficult bumps. Additionally, its performance is characterized by easily initiated and quickly executed turns. Nevertheless, exactly how fast the Queen can glide is undetermined as of now. But without question one of its distinguishing characteristics is a great deal of durability. In fact, if you ski fewer than ten times a season, it will probably have a life of five years; if you ski more than ten times a season, it will probably have a life of three years. The Queen can be used with most conventional bindings. However, it functions best with the Salomon Double. It is available in all the standard lengths, although it is easiest to obtain in the five to six foot range. Fortunately, it can be ordered in any of six different colors. Finally and most importantly, the Hart Queen costs only one hundred and twenty-five dollars. Thus it costs about thirty dollars less than skis of comparable quality.

Paragraph 3b (with its main topics underlined):

For beginning and intermediate skiers an excellent ski is the Hart Queen. A very thin layer of tempered ash, ash direct from the hardwood forests of Kentucky, is its core. And to provide for added strength and flexibility, two major innovations are involved in its outer construction. For increased strength, two sheets of ten-gauge steel are securely molded to its layer of ash. For increased flexibility, a wrapping of highly active fiberglass surrounds its two steel sheets. As a result, three to one is the Queen's impressive flexibility ratio. On difficult bumps, an even better ratio develops. Additionally, easily initiated and quickly executed turns characterize it. Nevertheless, it is undetermined as of now exactly how fast the Queen can glide. But without question a great deal of durability is one of its distinguishing characteristics. In fact, a life of five years for it is probable if you ski fewer than ten times a season; a life of three years for it is probable if you ski more than ten times a season. Most conventional bindings can be used with the Queen. However, the Salomon Double helps it function best. All the standard lengths are available for it, although the five to six foot range is the easiest to obtain. Fortunately, any of six different colors can

be ordered for it. Finally and most importantly, only one hundred and twenty-five dollars is the Hart Queen's price. Thus skis of comparable quality cost about thirty dollars more than it.